Biogeographic Research and the Spatial Web: Towards Interoperable Analysis Tools for Global Change Based on the Web Processing Service (WPS)

Christian GRAUL and Alexander ZIPF

Abstract

Global change scientists become increasingly aware of the importance of SDI, although current SDI are still limited in their analytical functionality. The recently introduced OGC Web Processing Service (WPS) specification allows for almost any GIS functionality to be built in SDI. Thus, it reveals new possibilities for further development of analysing and modelling data related to global change. Here we try to contribute in defining and implementing WPS in the domain of biogeographic research.

1 Introduction

GIS are an established tool for management, analysis and visualisation of spatial data in environmental research. Due to the onward development of open GI standards through the Open Geospatial Consortium (OGC), standardized web-based GIS services become available, accompanied by *Spatial Data Infrastructure* (SDI) development. As the access to spatial data of various domains and regions is of great significance for global change research, scientists become increasingly aware of the importance of SDI (LONGHORN 2003; e.g. VAN DER WEL 2005, BERNARD & OSTLÄNDER 2008). Along with data access, there is also an increasing demand for web-processing spatial data. Recently there has even started work on integrating OGC services for SDI in Grid-computing environments (e.g. LANIG et al. 2008).

Since global change research is a highly interdisciplinary field, standardization and interoperability are particularly crucial. Here SDI based on Service Oriented Architectures (SOA) seem to be an adequate approach. However, current SDI merely offer query, management and visualization of spatial data. Therefore the recently introduced OGC *Web Processing Service* (WPS) specification reveals new possibilities for further development of analysing and modelling data related to global change, such as biogeographic data.

Several projects (FÖRSTER & STOTER 2006, FRIIS-CHRISTENSEN et al. 2006, KIEHLE et al. 2006, DI et al. 2007, DIÀS et al. 2008, GERLACH et al. 2008) and our own work (STOLLBERG & ZIPF 2007, WEISER & ZIPF 2007, MAYER et al. 2008, STOLLBERG & ZIPF 2008) show that elaborate processing of spatial information using the WPS is feasible in several domains. However, there is still a lack of substantial functionalities, a common agreement on basic operations, a distinction between generic and domain specific functions or their sophisticated implementation, etc. (GÖBEL & ZIPF 2008). We try to contribute in defining and implementing those in the domain of biogeographic research similar to an example study conducted by GRAUL (2007), i.e. modelling the influence of climate change on vegetation.

2 Web-based geo-processing

Several fundamental geospatial OGC web standards (OWS) are now technically mature and widely used, namely *Web Feature Service* (WFS), *Web Coverage Service* (WCS), *Web Map Service* (WMS) and *Catalogue Service for Web* (CSW). However, these standards solely allow for retrieval and visualisation of geospatial data and therefore do not tap the full potential of SDI. Only the release of the WPS specification (SCHUT 2007) offers any sort of geospatial functionality to be built into SDI in a standardized way.

An OGC conform WPS must provide at least three mandatory operations, that allow for interoperability. The *GetCapabilities* operation returns service metadata, i.e. names and general descriptions of the processes. Detailed information about a single process, namely inputs, formats and outputs, is provided by the *DescribeProcess* operation. Finally the *Execute* operation runs a specified process. The WPS standard is quite unspecific, i.e. it may offer the subtraction of spatially referenced numbers, or a global climate change model (SCHUT 2007). It is obvious that the supply of single atomic web processes, like intersection or buffering, is only the starting point. Usually geospatial data analyses or modelling tasks consist of distinct process steps linked together. Users, humans as well as "machines", must be able to integrate single processes to carry out (semi-) automated high-order data processing, which is one key concept of future SDI (ALAMEH 2003). Certainly for reasons of efficient and elaborate implementation alone, it is necessary to define workflows through service chaining, using a modular set of atomic processes. Service chaining is usually termed *Web Service Orchestration* (WSO), but see PELZ (2003) to distinguish associated terms.

3 Concept of a biogeographic WPS

To examine a real life research workflow, we transfer a complex offline analysis, conducted with traditional desktop-GIS and further tools, to a web service rather than designing a new, probably "artificial", one. The original study (GRAUL 2007), addressing the impact of climate change on the invasibility (a region's susceptibility to plant invasions) in Germany, was not aimed at a web-based analysis.

The background of the study is the ongoing debate about the facilitation of biological invasions by climate change (e.g. DUKES & MOONEY 1999), an important aspect for invasion management and nature conservation. The object of the study was the determination of the influence of climatic parameters on the invasibility of German landscapes. Based on identified relationships between the number of non-native species and various environmental factors (e.g. geology, land-cover, DEM, vectors of dispersal), a grid-based statistical model was built. Subsequently the potential future invasibility due to climate change was predicted using wide range climate scenarios. A conceptual workflow is given in Fig. 1.

The development of the prototypic biogeographic WPS, based on the study briefly described above, is to be carried out in the following three main steps:

1. Transfer of the exemplary study to a chained WPS analysis
2. Develop a client for interactive workflow design
3. Open the base WPS with the client for external services

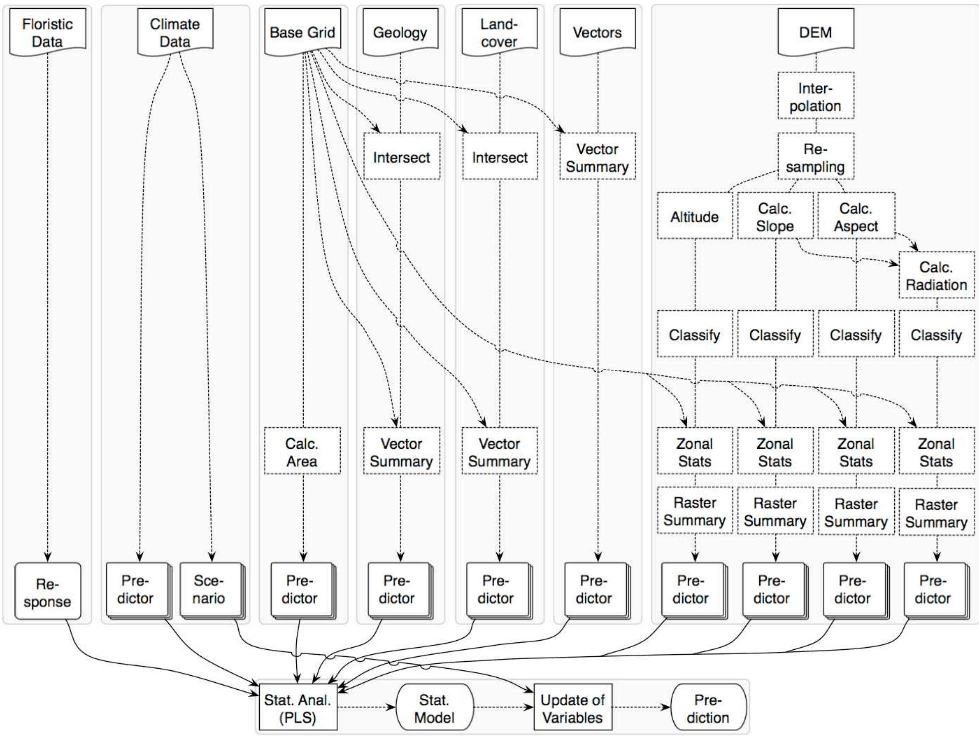

Fig. 1: Conceptual workflow of the original study (after Graul 2007)

These three steps raise a couple of corresponding questions, some of which are particularly apparent. As to the first step it needs to be clarified which atomic processes are required to fulfil the desired functionality, keeping the possibility of general and flexible use in mind. Furthermore techniques to incorporate elaborate statistical methods have to be examined. As there are particular requirements (see below), special attention should be drawn to the implementation of the service chaining.

Questions related to the second step are slightly different in nature. Although the technical background is also certainly important, usability plays a key role here. Workflows should not be fixed but freely designable by users. Hence they have to be provided with a sophisticated full-featured client, without compromising the straightforward character of web-based geo-processing. The most appropriate user interface in this case is certainly a flexible graphical workflow builder (GWB). Very first steps in this direction are presented by FÖRSTER & SCHÄFFER (2007). Sometimes user intervention is required between the process steps, e.g. to compare two intermediate outputs visually and decide which one to use in the further processing. Therefore the service should be able to handle such interaction. This, in turn, affects the implementation of both, the GWB and the orchestration.

Finally the question to be addressed in the third step is how to best integrate external data and services, thus realizing the basic idea of SDI – the consolidation of decentralized spatial data and services from several providers. Both of them must be attachable in the client, enabling the user to process extensive data sources using a broad range of operations.

4 Implementation

Working on the first step, we use deegree2 as underlying WPS framework (FITZKE et al. 2004). Analysing the original study's workflow yields in 11 core GIS processes that need to be transferred to the WPS (Table 1). These are implemented in Java, using especially the GeoTools, JTS and JAI libraries. For grid data processing BAUMANN (2008) recently introduced a WCS processing extension (WCPS), which is now a candidate OGC standard. Its capabilities, especially as to WSO, need to be tested and compared to WPS. Two statistical processes were identified so far. However, this list is highly extendable, thus it seems appropriate to delegate such tasks to an existing environment. As a widely used open-source statistical environment, the R-project is most applicable here. Rserve is a straightforward way to provide the full variety of R functionality to applications (URBANEK 2003). There is even a Python-based WPS available, that encapsulates R (CEPICKY 2006), but we chose to implement the identified statistical processes through the Java interface.

WSO is carried out by an *Orchestration Engine* (OE), which coordinates the interacting services involved, controlled by a document containing chaining instructions in a certain description language. The *Business Process Execution Language* (BPEL) provides a de-facto WSO standard. However, WEISER & ZIPF (2007) as well as STOLLBERG & ZIPF (2007) pointed out several problems in the use of BPEL together with current OWS versions, concerning e.g. communication protocols and the transfer of raw binary data. Thus, the latter suggested the use of the WPS interface itself to chain services. In this case one service acts as OE, coordinating the others. As there is no satisfying alternative approach, we continue this work, keeping track of the current OGC activity on this topic. Most recently a workflow working group was formed and we expect a sophisticated solution in the near future.

Table 1: Identified processes to be transferred to the WPS (green = GIS, blue = Statistics)

Process	**Description**	**Status**
Area	Calculates the polygon's area of a vector layer	operational
Intersect	Intersects two vector layers	operational
ZonalVectorSummary	Calculates area ratios of polygons or length of lines of one layer in zones given by a second layer	planed
Interpolation	Interpolates a grid coverage from points	planed
Resampling	Resamples a grid coverage	in progress
Slope	Calculates slope using a DEM grid	operational
Aspect	Calculates aspect using a DEM grid	operational
SolarRadiation	Calculates radiation using slope and aspect grids	operational
Reclassification	Reclassifies grid coverage values	operational
ZonalStatistics	Calculates grid statistics in a vector zone layer	operational
ZonalRasterSummary	Calculates area ratio of a classified grid layer in zones given by a vector layer	planed
PLSRegression	Processes a PLS-regression analysis	in progress
Prediction	Calc. predictions using a model and scenario	in progress

5 Conclusions and outlook

Like cooperation is all-important in global change research, interoperability is on the technical level. Thus, we need standards that allow for technical compatibility to, in turn, achieve better scientific cooperation. Based on the WPS specification, we are able to implement various geo-processes and statistical processes as standardized web-services. Moreover we can use a WPS as a kind of OE to aggregate several processes to complex workflows. Finishing the redesign of the original study by GRAUL (2007), based on OpenGIS standards, lays the fundament for further studies, with similar data or questions that utilize these functions. It will be available in a service-based environment, so that the services can be integrated in web-based or even desktop applications. This allows defining new service chains and plug-in new processes in a modular and easy way, without changing existing applications or knowing implementation details, language or platform of the services. Apart from well-known basic GIS functions a suite of higher level functionalities, more specific for this kind of studies, can occur, that is build on top (or aside) these basic GIS and statistical operations. We see this work as a first step in providing processes for a domain-specific WPS specification profile as well as a step towards a richer toolbox of GIS functions for global change research based on the modern paradigm of SOA.

References

ALAMEH, N. (2003), Chaining geographic information web services. IEEE Internet Computing 7: 22–29.

PELZ, C. (2003), Web services orchestration and choreography. Computer 36: 46–52.

BAUMANN, P. (2008), Web Coverage Service (WCS) – processing extension. Version 0.0.9, Candidate OpenGIS Implementation Standard 08-059r2.

BERNARD, L., OSTLÄNDER, N. (2008), Assessing climate change vulnerability in the arctic using geographic information services in spatial data infrastructures. Climatic Change 87: 263–281.

CEPICKY, J. (2006), GRASS goes web: PyWPS. Free and Open Source Software for Geoinformatics (FOSS4G), Lausanne, Switzerland.

DI, L., ZHAO, P., HAN, W., WEI, Y., LI, X. (2007), Web service-based GeoBrain online analysis system (GeOnAS). NASA Science Technology Conference, Adelphi, USA.

DIÀZ, L., GRANELL, C., GOULD, M. (2008), Case study: geospatial processing services for web-based hydrological applications. In: J. T. Sample, K. Shaw, S. Tu, M. Abdelguerfi: Geospatial Services and Applications for the Internet. Springer.

DUKES J. S., MOONEY H. A. (1999), Does global change increase the success of biological invaders? Trends in Ecology and Evolution 14: 135–139

FITZKE, J., GREVE, K., MÜLLER, M., POTH, A. (2004), Building SDIs with free software – the deegree project. In: Proceedings of GSDI-7, Bangalore.

FÖRSTER, T., SCHÄFFER, B. (2007), A client for distributed geo-processing on the web. In: J. M. Ware, G. E. Taylor: W2GIS 2007, Lecture Notes in Computer Science 4857: 252–263

FÖRSTER, T., STOTER, J. (2006), Establishing an OGC web processing service for generalization processes. Workshop of the ICA Commission on Map Generalisation and Multiple Representation, Portland, USA.

FRIIS-CHRISTENSEN, A., BERNARD, L., KANELLOPOULOS, I., NOGUERAS-ISO, J., PEEDELL, S., SCHADE, S., THORNE, C. (2006), Building service oriented applications on top of a spatial data infrastructure – a forest fire assessment example. 9th AGILE Conference on Geographic Information Science, Visegrád, Hungary.

GERLACH, R., SCHMULLIUS, C., NATIVI, S. (2008), Establishing a web processing service for online analysis of earth observation time series data. Geophysical Research Abstracts 10.

GÖBEL, R., ZIPF, A. (2008), How to define 3D geoprocessing operations for the OGC web processing service (WPS)? Towards a classification of 3D operations. International Workshop on Computational Geoinformatics (CompGeo), Perugia, Italy.

GRAUL, C. (2007), Modellierung der Invasibilität deutscher Landschaften unter dem Einfluss des Klimawandels. Diploma Thesis, University of Bayreuth.

KIEHLE, C., GREVE, K., HEIER, C. (2006), Standardized geoprocessing – taking spatial data infrastructures one step further. 9th AGILE Conference on Geographic Information Science, Visegrád, Hungary.

LANIG, S., SCHILLING, A., STOLLBERG, B., ZIPF, A. (2008). Towards standards-based processing of digital elevation models for grid computing through web processing service (WPS). International Conference on Computational Science and its Applications (ICCSA), Perugia, Italy.

LONGHORN, R. A. (2003), A comparison of spatial information access policies of transnational environmental modeling and global climate change programs. In: Benes, T.: Geoinformation for European-wide Integration. Millpress, Rotterdam: 305–313.

MAYER, C., STOLLBERG, B., ZIPF, A. (2008 submitted), Providing near real-time traffic information within spatial data infrastructures.

SCHUT, P. (2007), OpenGIS Web Processing Service. Version 1.0.0, OpenGIS Standard 05-007r7.

STOLLBERG, B., ZIPF, A. (2007), OGC web processing service interface for web service orchestration – aggregating geo-processing services in a bomb threat scenario. Web & Wireless GIS Conference (W2GIS), Cardiff, UK.

STOLLBERG, B., ZIPF, A. (2008), Geoprocessing services for spatial decision support in the domain of housing market analyses – experiences from applying the OGC web processing service interface in practice. 11th AGILE Conference on Geographic Information Science, Girona, Spain.

URBANEK, S. (2003), RServe: a fast way to provide R functionality to applications. Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria.

VAN DER WEL, F. J. M. (2005), Spatial data infrastructure for meteorological and climatic data. Meteorological Applications 12: 7–8.

WEISER, A., ZIPF, A. (2007), Web service orchestration of OGC web services for disaster management. 3rd International Symposium on Geoinformation for Disaster Management (GI4DM), Toronto, Canada.